Evaluating Changes in Cognitive Processing During Skill Development: A Structural Equation Modeling Approach

David J. Law

Tennessee Department of Health

Paper Presented at the Annual Meeting of the American Educational Research Association
Seattle, Washington
April 13, 2001

Address: Cordell Hull Building, 6th Floor
425 5th Avenue North
Nashville, TN 37247
dlaw@mail.state.tn.us

Abstract

A latent growth curve analysis evaluating the dynamics of cognitive processing in developing skill is presented as an example of a methodological approach that combines structural equation modeling and change for the purpose of evaluating cognitive theory. Theory-based predictions of the relationships of working memory capacity and procedural memory speed with developing skill were tested in a series of latent growth curve analyses. Overall, results supported theoretical accounts of working memory as a major determinant of early performance in complex tasks, with a diminished working memory contribution and an increased procedural memory contribution following consistent practice. Confirmatory factor analyses of performance in verbal and visuospatial working memory tasks supported a domain-based correlated two-factor structure. Latent growth curve analyses of repetition priming showed that repetition priming is a multi-episodic process that takes the form of a log-log linear function, which was consistent with the hypothesis that repetition priming is isomorphic with the procedural memory processes underlying skill. Factor scores generated from these analyses were then used as indicators of working memory capacity and procedural memory speed in order to predict skilled performance. Skill development was measured as performance in a computer-based instructional environment that taught participants to compute the outputs of digital logic gate circuits. Baseline latent growth models showed that learning in the logic gate task took the form of a power function and that the levels of heterogeneity in performance levels and learning rates warranted attempts to create predictor models using the working and procedural memory indicators. The final predictor models showed that both working memory capacity and procedural memory speed affected learning rates. Furthermore, initial working memory contributions to performance were large while initial procedural memory contributions were small, and these relationships were reversed in later trials. Although the initial working memory factor analyses showed a two-factor solution, the portion of working memory variance predicting learning was the common variance among the two factors. Altogether, the results are consistent with theoretical accounts of working memory, procedural memory, and skill, and of the use of repetition priming levels as an indicator of procedural memory speed. The results also support claims that the functional aspect of working memory is a single general ability component. Finally, it is argued that the combination of strong theoretical predictions of change and structural equation modeling techniques for assessing change provides a powerful methodology for evaluating contemporary cognitive theories.

Assessing growth or change and its mediating factors is central to many of the important questions in contemporary educational research. Historically, the measurement of change was so problematic that it caused leading researchers to question whether we should even try (Cronbach & Furby, 1970). More recently, researchers have shown that the combination of multiple waves of data (Willett, 1988, 1989) and multilevel modeling techniques, such as hierarchical linear models (Bryk & Raudenbush, 1992) and latent growth curve models (Chou, Bentler, & Pentz, 1998; McCardle & Epstein, 1987; Willett & Sayer, 1994), is adequate for modeling both interindividual and intraindividual change. Accordingly, these methods are increasingly represented in educational research. However, most research using these methods has focused on macro-level units of analyses such as demographic groups, classrooms, and school systems over extended periods of time. In contrast, the present study presents a micro-level analysis of the dynamic performance-ability relationships that occur during cognitive skill acquisition. In many respects, this research is similar to cognitive correlates and cognitive components approaches, which were popular in the 70's and 80's (see, Pellegrino & Glaser, 1979), but with an updated methodology focused on change. The primary purpose of this paper is to show the utility of such an approach rather than reporting the substantive results, which are reported elsewhere (Law, 2000).

Background

The study used latent growth curve models to evaluate the changing relationships of working memory capacity and procedural memory speed with performance in a complex cognitive task. Latent growth curve analyses were also used to evaluate the validity of repetition priming as an indicator of procedural memory speed. This required a strong a priori model of skill learning and individual differences. The model combined significant aspects of Anderson's (1983, 1993) ACT-R, Logan's (1988) instance-learning, Ackerman's (1987, 1988) cognitive/psychometric, and Cohen's (1984, Cohen & Squire, 1980; Poldrack, Selco, Field, & Cohen, 1999) procedural memory/repetition priming models. In brief, initial performance in complex tasks is attention demanding and heavily dependent upon working memory capacity; and with practice, performance of the consistent task components becomes increasingly automatic - the product of a capacity-free procedural memory system. Accordingly, it was hypothesized that individual differences in the constructs representing working memory capacity and procedural memory speed will show systematically decreasing and increasing relationships with skilled performance over the course of its development. To test these hypotheses it was necessary to develop an individual differences marker for procedural memory.

Cohen and Squire (1980) were the first to suggest that the processes underlying repetition priming and procedural memory are one and the same. However, many researchers have come to accept as fact, findings that repetition priming is a single-episode phenomenon (e.g., Schacter, Cooper, Delaney, Peterson, & Tharan, 1991). If true, repetition priming could not underlie the gradual multi-episode process of skill learning. A major contention of this paper is that the two-to three-wave mean difference analyses of change, which were used in these studies, are insufficient to test the hypothesis of gradual change, and that multi-wave growth curve analyses provide a more appropriate test.

Methods and Procedures

Participants were 120 volunteers, a small but adequate number to test the proposed models. They were solicited from the Vanderbilt University community and paid $35 for their efforts, which required them to attend sessions on four consecutive days. The first and second sessions involved participating in separate verbal and visuospatial working memory and repetition priming tasks. On the third day participants were randomly assigned to either part or whole training conditions in computing digital logic gates. In the final session, all participants computed three-gate digital logic circuits under conditions identical to the whole-training condition. This paper reports on the results from the first three sessions.

Working memory was assessed using tasks that required the concurrent processing and storage of information (e.g., Baddeley, 1986). Repetition priming was assessed using a lexical decision task and a symmetry decision task (for details see, Law, 2000). Repetition priming is the phenomenon such that processing of a stimulus is enhanced in presentations following the initial presentation of a stimulus. The enhancement is shown by increased accuracy and decreased latency of processing in subsequent presentations. The repetition priming hypotheses evaluated in the study were whether the priming effect continues to increase with subsequent priming episodes, and if so if priming takes the form of a power or a log-log linear function, the form commonly taken by developing skill (e.g., Newell & Rosenbloom, 1981). A positive finding for both of these hypotheses was a prerequisite for using repetition priming as an indicator of procedural memory speed in the cognitive skill development models.

Cognitive skill development was assessed by monitoring participants' performance as they learned to compute the output of digital logic gates, either as single-gates or three-gate circuits. Examples of the single-gate and three-gate stimuli are presented in Figure 1. The single gate example is an OR gate and the circuit example is a NAND gate and an XOR gate feeding into an NXOR gate. Digital logic gates process digital input, zeros and ones, compute a logical function, and produce digital output dependent upon their input and function. Thus, participants were required to learn the logic functions associated with each of six different gate types, match the functions to their schematic representations, and compute their outputs for various sets of inputs.

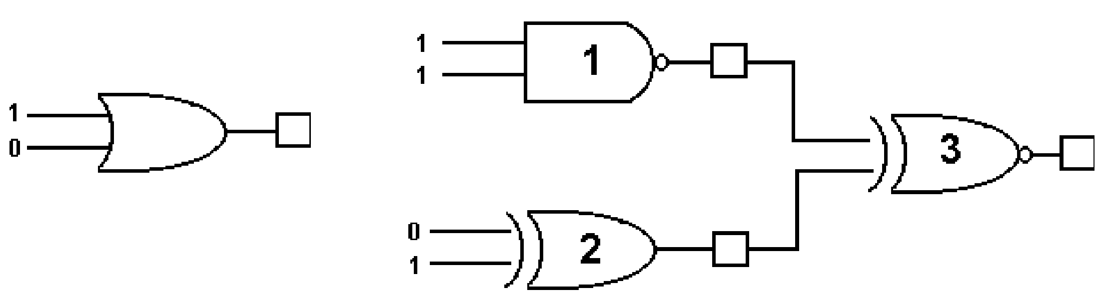

Figure 1. **Single- and multiple-gate digital logic stimuli.**

All of the tasks were computer-administered. Span length and the accuracy of concurrent processing were recorded in the working memory tasks, and accuracy and latency in milliseconds were recorded in the repetition priming and logic-gate training tasks. The primary measures of interest were span length in the working memory tasks and latency in the repetition priming and logic-gate training tasks. However, using these data, span length and latency, in the modeling procedures required that the corresponding accuracy levels were high and there were no systematic differences between the two training groups. These conditions were met and the analyses proceeded.

Results and Discussion

The working memory span-length data were evaluated using a series of alternate confirmatory factor models. Ultimately, only a correlated two-factor solution with factors representing the verbal and visuospatial domains fully accounted for the data. Accordingly, factor scores were derived from the correlated two-factor solution and used as predictors in the skill learning growth curve models.

Before proceeding to the latent growth curve models it is worthwhile to look at individual growth and consider how it was assessed in the priming tasks. The models were assessed in a log-log space; that is, log latency was modeled in terms of log episode. This was done with the expectation of priming taking the form of a log-log linear function, which is an alternate form of a power function. Figure 2 shows the individual log-log linear least-squares trajectories for twelve randomly selected participants in the verbal priming task (left panel) and visuospatial priming task (right panel). Inspecting Figure 2, it is apparent that there was considerable variation in the individual functions. On average the intercepts and slopes were larger in the visuospatial condition than in the verbal condition. There was also greater heterogeneity among the intercepts and slopes in the visuospatial condition. These observations suggest the participants were more alike and more facile in dealing with verbal stimuli than visuospatial stimuli. Taken in context these findings seem reasonable; that is, given the participants were from a college educated population and the stimuli in the lexical decision task were common five-letter words, while the stimuli in the symmetry decision task were irregular polygons created specifically for this study.

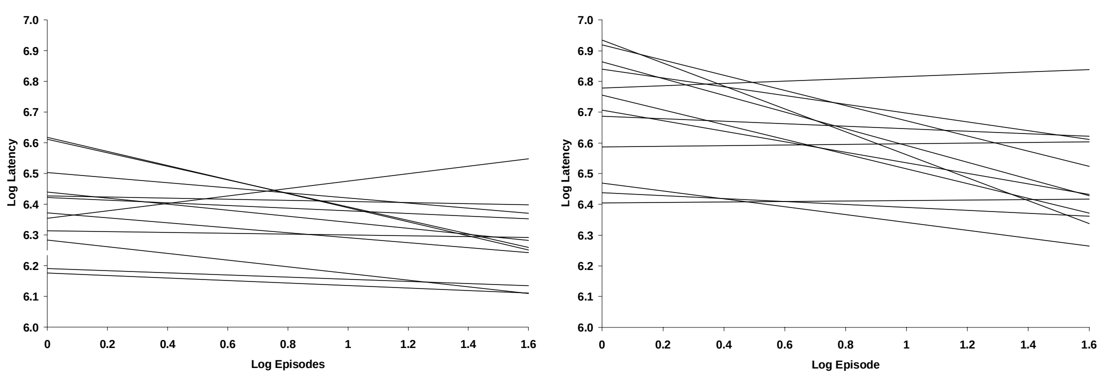

Figure 2. **Individual log-log linear functions describing verbal (left panel) and visuospatial (right panel) priming for 12 randomly selected participants.**

Growth-Curve Factor Structures. The factor structure presented in Figure 3 represents a five-episode latent growth model. In the model there are five observed variables, T1 through T5, which represent five consecutive priming episodes measured in log units. There are also two latent variables, INT and SLP, connected to the observed variables via factor loadings. The latent variables represent the true intercept and true slope for the population function describing change, and it is the means and variances of these variables that are the parameters of primary interest in this model. Generally factor loadings are variables and the estimated values are in large part the answer to the questions posed by the model. In this case, however, the factor loadings are set to constant values. Each of the INT factor loadings are set to 1 and the SLP factor loadings are set to the log values of the linear constants 1 through 5. Combining these loadings with log-unit observed data forces the model to estimate a five-episode log-log linear function. The two-headed arrow between INT and SLP represents the covariance of the true intercept and true slope. These are estimated as part of the modeling process. Finally, the arrows below the observed variables represent error variances corresponding to their measurement error. These too, are estimated as part of the modeling process.

Indicators of overall model-fit are then used to determine whether the observed data are consistent with the hypothesized model. Given an adequate overall model-fit, the parameter estimates can provide useful information for describing the process of change. The mean estimates of true intercept and true slope describe the overall trajectory of growth in the population and the variance estimates indicate the degree of individual differences in both the level of performance and the rate of growth. The covariance of the true intercept and true slope indicates whether the level and rate of growth in the population are correlated. Additionally, the error variances can be used to estimate the reliability of measurement for both the observed variables and the overall rate of growth. In the repetition priming analyses, factor scores representing the true intercepts and slopes were generated from the latent growth curve functions. Subsequently, these scores were used to predict the true intercept and true slope in the logic-gate training task growth models.

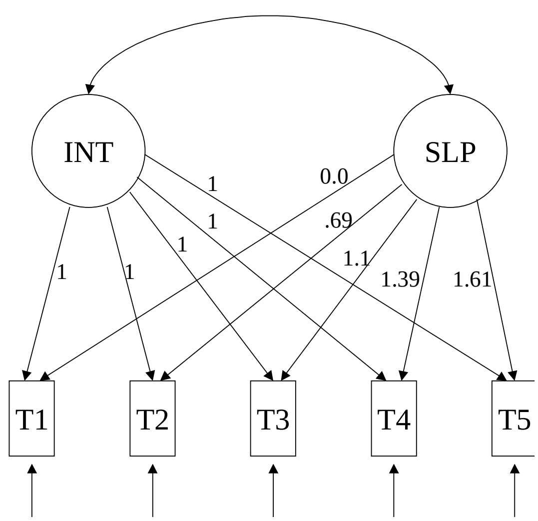

Figure 3. **Schematic representation of a five episode latent growth curve model.**

Assessing Model-Fit. The distributions of all of the data modeled in the growth curve analyses deviated from multivariate normal. This is not unusual for this type of performance data. However, data that is not multivariate normal can cause problems in the estimation and assessment of covariance structure models, such as the latent growth curve models used in this study. Furthermore, the degree of difficulty encountered will likely be proportional to the degree of deviation. Data that deviates from multivariate normal, and especially kurtotic data, has been shown to cause increased chi-square statistics and decreased goodness-of-fit indicators, both of which can cause the rejection of a model that should be accepted (Hu, Bentler, & Kano, 1992; Hu & Bentler, 1995). Multivariate kurtosis can also cause standard error estimates for individual parameter estimates to be biased (Chou & Bentler, 1995).

However, researchers have developed robust methods to deal with these problems. The Satorra-Bentler Scaled X^2 statistic and the CFI goodness-of-fit index have been shown to be robust to deviations from multivariate normality and to perform well in small to moderate sized samples (Hu, et al. 1992; Chou & Bentler, 1995; West, Finch, & Curran 1995). Robust standard error estimates have also been developed that provide unbiased significance tests under nonnormality (Arminger & Schoenberg, 1989). Accordingly, the overall growth curve models were assessed using the Satorra-Bentler Scaled X^2 statistic in conjunction with the CFI. Additionally, the statistical significance of the individual parameter estimates was evaluated using robust standard errors. The standardized root-mean square residual (SRMR), a chi-square statistic computed under an assumption of normality, and the GFI goodness-of-fit index are also provided with each model for comparative and informational purposes. The standard X^2 and Scaled X^2 statistics both assess the lack of fit due to model-based constraints. The CFI, an incremental fit index, indexes the relative reduction in the lack of fit by the specified model versus a baseline model, whereas the GFI, an absolute fit index, indexes the relative amount of variances and covariances accounted for by the specified model. Under nonnormality the standard X^2 is expected to be inflated and the GFI deflated relative to their 'true' values. Finally, the SRMR estimates the average absolute value of the discrepancy between observed and predicted correlations among the observed variables.

Repetition Priming Models. The parameter estimates and model-fit statistics from the latent growth curve analyses of visuospatial and verbal repetition priming are presented in Table 1, the parameter estimates in rows 1 through 10 and the model-fit statistics in rows 11 through 14. There are two latent growth curve models for visuospatial priming. SPATIAL-1 is a five-episode log-log linear model assuming homoscedastic error, and SPATIAL-2 assumes heteroscedastic error. Likewise there are two five-episode verbal priming models, VERBAL-1, a homoscedastic model, and VERBAL-2, a heteroscedastic model.

Looking at the model-fit statistics for the visuospatial models, SPATIAL-1 appears to provide an adequate description of the visuospatial priming data. Though the Scaled X^2 for SPATIAL-1 was slightly large for its degrees of freedom, X^2 (14) = 28.89, both of the goodness-of-fit indicators, GFI = .924 and CFI = .971, were larger than the customary criterion of .90. Additionally, the standardized root mean-square residual, SRMR = .055, was small relative to the magnitude of the correlations that were observed among these variables. Altogether, this suggests that SPATIAL-1 adequately described the changes in latency that occurred over the course of five visuospatial priming episodes. Finally, the chi-square difference test used to compare Models SPATIAL-1 and SPATIAL-2, X^2 (4) = 5.57, p = .23, showed that allowing the error variances to vary did not add to the overall model-fit anymore than would be expected to occur by chance. Accordingly, SPATIAL-1 was the preferred model of repetition priming with visuospatial stimuli.

Table 1 ***Repetition Priming: Log-log Linear Latent Growth Curve Models***

	Maximum Likelihood Estimates			
Parameter	SPATIAL-1	SPATIAL-2	VERBAL-1	VERBAL-2
1. Intercept Mean	6.6844***	6.6837***	6.3922***	6.3917***
2. Slope Mean	-.1172***	-.1168***	-.0984***	-.0971***
3. Intercept Variance	.0619***	.0629***	.0299***	.0287***
4. Slope Variance	.0123***	.0135***	.0060***	.0054***
5. INT/SLP Covariance	-.0162***	-.0172***	-.0100***	-.0092***
6. Error 1	.0096***	.0063	.0040***	.0057**
7. Error 2	.0096***	.0093***	.0040***	.0054***
8. Error 3	.0096***	.0121***	.0040***	.0038***
9. Error 4	.0096***	.0069***	.0040***	.0035**
10. Error 5	.0096***	.0109***	.0040***	.0028**
11. (df) X^2	(14) 30.10	(10) 23.99	(14) 53.45	(10) 45.18
12. Scaled X^2	(14) 28.89	(10) 22.32	(14) 43.04	(10) 38.90
12. GFI	.924	.945	.893	.912
13. CFI	.971	.975	.925	.934
14. SRMR	.055	.053	.060	.049

Note: $^{**}p < .01$; $^{***}p < .001$; N=119; INT= Intercept; SLP = Slope.

All of the individual parameter estimates in SPATIAL-1 were statistically significant (p's < .001). Entries in the first two rows of Table 1 are estimates of the population means of true intercept, 6.684, SE = .0242, and true slope, -.1172, SE = .0124. Together, these describe the trajectory of change that was visuospatial repetition priming. Entries in the third through fifth rows are the estimated population variances of the true slope and true intercept and their covariance, .0619, SE = .0078, .0123, SE = .0025, and -.0162, SE = .0034, respectively. That the intercept and slope variances were statistically significant shows that there was indeed interindividual heterogeneity in both the participants' initial status and rate of change. While this is of interest in itself, it was also an essential precondition for the slope and intercept factor scores being used as predictors of learning in the logic-gate training task. That the covariance of the slope and intercept estimates was statistically significant shows that the participants' initial status and rate of change were related, $r(117) = -.587$, $p < .001$. Keeping in mind that the true slope was negative, this indicates that the individuals who were initially highest in log-latency (i.e., slowest) showed the greatest rate of change. Although this may be a simple artifact of extreme scores and regression to the mean, it is also consistent with an account such that slower individuals have fewer relevant memory traces and are thus on an earlier and faster declining portion of the learning curve.

Model-based reliability estimates for the five priming episodes ranged from, r = .81 to r = .87. Thus, task performance was measured with a high degree of reliability. Finally, the estimated reliability of change in log-latency was remarkably high, r = .93, that is, given the controversy that has often surrounded the measurement of change (e.g., Cronbach & Furby, 1970; Willett, 1988). Altogether this model shows that the decreasing latency that is repetition priming is not a single episodic event. Rather, repetition priming is a sustained process that can continue over five or more episodes and in doing so takes the form of a decreasing log-log linear function.

Turning to the verbal-priming models, the Scaled X^2 for VERBAL-1 was somewhat large for its degrees of freedom, X^2 (14) = 43.04, as was the Scaled X^2 for VERBAL-2, X^2 (10) = 38.90. However, the CFI goodness-of-fit indicators for both VERBAL-1 and VERBAL-2 were acceptable, CFI = .925 and CFI = .934. The SRMR was also relatively small for both models, SRMR = .060 and SRMR = .049. Given the extreme kurtotic nature of these data, the CFI and SRMR values suggest both models were acceptable. That the chi-square difference test was not statistically significant, X^2 (4) = 4.14, p = .39, indicates VERBAL-1 was preferred.

As in the visuospatial priming model, all of the parameter estimates in VERBAL-1 were statistically significant (p's < .001). The true intercept, 6.392, *SE* =.0167, and true slope, -.0984, *SE* = .0085, described a decreasing function with significant heterogeneity in both initial performance level, .0299, *SE* = .0039, and rate of change, .0060, *SE* = .0010. Also similar to visuospatial priming, the initial performance level and the rate of change were correlated, $r(117) = -.751$, $p < .001$. Reliability estimates for the individual episodes ranged from, r = .76 to r = .88, and the rate of change was measured with a reliability of, r = .94. Altogether, the verbal and visuospatial priming results show that repetition priming is a continuously decreasing process that spans five or more episodes. Furthermore, the change that is priming can be measured with high degree of reliability and takes the form of a log-log linear function. Accordingly, model-based factor scores corresponding to the true intercept and true slope in the visuospatial and verbal priming models were generated for each participant and used as predictors of performance in the logic-gate training task.

Baseline Logic-Gate Models. The first step in assessing the individual growth hypotheses was to create baseline growth models, evaluating both the form of the learning curves and the degree of individual differences in the model parameters. As in the priming models, the hypothesized form was log-log linear and the parameters of interest were the population slopes and intercepts, which corresponded to the population learning rates and performance levels, respectively. Though not as extreme as in the priming data, the logic-gate data also deviated from multivariate normality. Accordingly, model-fit evaluations were based primarily on the Scaled X^2 statistic, CFI, and SRMR.

Two baseline models were evaluated, one for the positive gates and one for the negative gates, both with homoscedastic error structures. The parameter estimates and model-fit statistics for the baseline models are presented in Table 2. The parameter estimates are in rows 1 through 6 and the model-fit statistics in rows 7 through 10. The apparent discrepancies between the standard minimum-fit and Scaled X^2 statistics and the GFI and CFI indices reflect the deviation from multivariate normal. Regardless, the

Scaled X^2 statistics were small relative to their degrees of freedom, X^2 (38) = 47.54, and X^2 (38) = 47.81, the CFI were large, CFI = .971 and CFI = .981, and the SRMR were small, SRMR = .040 and SRMR = .035, in the positive and negative gates, respectively. Additionally, all of the parameter estimates except for the intercept/slope covariances were statistically significant (p's < .001). Thus, the log-log linear models provided adequate descriptions of the logic gate training data and there were sufficient individual differences in the intercept and slope estimates to warrant further investigation.

There were also a number of interesting differences among the positive and negative gates. The mean intercept was lower, 7.7343, *SE* = .0281, and the mean slope steeper, -.2268, *SE* = .0091, in the positive gates than in the negative gates, 8.0319, *SE* = .0232 and -.1938, *SE* = .0077, (Based on an interval of ± 2 *SE*). This indicates that participants were initially faster and acquired skill at a greater rate in the positive gates. There were also greater individual differences in the positive intercept, .0860, *SE* = .0110, than in the negative intercept, .0581, *SE* = .0079. However, the difference in slope variances was not statistically significant, .0065, *SE* = .0015, and .0045, *SE* = .0010. The estimated reliabilities for the individual blocks ranged form r = .88 to r = .89 in the positive gates and from r = .87 to r = .90 in the negative gates. Thus, the logic-gate training task provided reliable measurement in both the positive and negative gates. Finally, the estimated reliabilities for the true slopes were high, r = .996, for both gate types. Such a high degree of reliability reflects the efficiency of using eight waves of data to model change (see, Willett, 1989).

Table 2 ***Eight-Block Log-Log Linear Logic Gate Baseline Models***

	Maximum Likelihood Estimates	
Parameter	Positive Gates	Negative Gates
1. Intercept Mean	7.7343***	8.0319***
2. Slope Mean	-.2268***	-.1938***
3. Intercept Variance	.0860***	.0581***
4. Slope Variance	.0065***	.0045***
5. INT/SLP Covariance	-.0053†	.0002
6. Error 1-8	.0110***	.0087***
7. (df) X^2	(38)81.73	(38)66.25
8. Scaled X^2	(38)47.54	(38)47.81
9. GFI	.868	.889
10. CFI	.971	.981
11. SRMR	.040	.035

Note: †p < .10; ***p < .001; N=119; INT = Intercept; SLP = Slope.

Together, the two baseline models provided an informative description of learning in the logic gate task. That the form of the learning curve was log-log linear fulfilled a strong expectation based on the log-log linear law of learning (Newell & Rosenbloom, 1981). This is important because fulfilling the expectation supports the validity of the measurement models that form the bases of these analyses. There was also a high degree of individual differences in both the level of performance and rates of change. Additionally, systematic differences associated with differences in the positive and negative gates were seen in both the performance levels and rates of change. Overall performance was higher and the rate of change greater in the positive gates than in the negative gates. Likewise, the magnitude of individual differences was higher in the positive gate condition. Both of these findings reflect the added complexity of the negative gate judgments. The predictor models that follow will determine to what extent these differences can be accounted for by individual differences in capacity, speed, and part/whole-training.

Logic Gate Models with Predictors, The predictor models represent the simultaneous regressions of the true slopes and true intercepts on the predictor variables. Thus, each coefficient estimates the change in the true slope or true intercept that is uniquely attributable to a unit change in a given predictor. In this context, creating and evaluating the predictor models was an iterative process of simultaneously evaluating overall model-fit and seeking the most informative and parsimonious models. Initially, models were assessed that included all of the predictor variables and those predictors that failed to show a significant relationship with either the logic-gate slope or intercept were dropped from subsequent analyses. The entire process was resolved in two iterations, the initial models and the final models. In both iterations two models were assessed for both the positive and negative gates, one that fit the intercept as initial performance level and one that fit the intercept as final performance level. Together, these models provided estimates of initial and final performance levels, the rate of change over the course of training, and the mediating effects of training method, procedural speed, and working memory capacity on each of these parameters.

To fully understand the nature of the relationships that were obtained in the predictor models, it is useful to examine the simple correlations of the predictors with the true slopes and intercepts. These correlations were derived from the LISREL ETA and KSI matrices that came from the initial predictor models and are presented in Table 3. The verbal and spatial PS correlations were similar across the positive and negative gate conditions, with one difference. Verbal PS was significantly correlated with the positive-gate learning rate and spatial PS was not, whereas spatial PS was significantly correlated with the negative-gate learning rate and verbal PS was not. Otherwise, neither PS level indicator correlated with initial positive-gate performance, both were correlated with positive-gate final performance, and both correlated with negative-gate initial and final performance levels. However, when simultaneously regressed in the initial models, verbal PS was the only significant PS indicator in the positive-gate condition and spatial PS was the only significant PS indicator in the negative-gate condition.

Table 3 ***Correlations of Predictors with Latent Curve Slopes and Intercepts***

Predictor/Parameter	Early Positive	Late Positive	Early Negative	Late Negative
1. SP PS /INT	.159	.248**	.245**	.314***
2. SP PS/SLP	.170	--	.208*	--
3. VB PS/INT	.146	.256**	.184*	.232*
4. VB PS/SLP	.207*	--	.149	--
5. SP WM/INT	-.307***	-.200*	-.262**	-.202*
6. SP WM/SLP	.175*	--	.047	--
7. VB WM/INT	-.364***	-.244**	-.332***	-.264**
8. VB WM/SLP	.195*	--	.043	--
9. Group/INT	.292**	.528***	.323***	.512***
10. Group/SLP	.443***	--	.468***	--

Note: **p* < .05; ***p* < .01; ****p* < .001; N = 119; Note: **p* < .05; ***p* < .01; ****p* < .001; SP = Spatial; VB = Verbal; PS = Procedural Speed Level; WM = Working Memory; INT = Intercept; SLP = Slope.

Given the simple correlations, which showed both PS indicators significantly correlated with performance in both the positive and negative gates, these findings show the combined effects of multicollinearity and the differential weightings of the PS variables in the positive and negative gates. Though these differences may simply be artifactual, the differences in weightings across gate types and verbal and visuospatial priming suggest possible differences in processing. It may be that the positive-gates were a more straightforward verbal/logical task, requiring verbal mediation, whereas the negative gates with the added visual feature of the negation node (see, Figure 1) and the possible use of an added representational dimension for negation, caused participants to rely on visuospatial representations. This would be similar to performance differences that have been observed in simple digit-span tasks where the digit-span forward is seen as the product of simple verbal attention and the backwards digit-span is associated with visual imagery (Weinberg, Diller, Gerstman, & Schulman, 1972). Overall, this supports some degree of specialization within the PS domains.

The simple correlations of the verbal and spatial WM variables were almost identical across the positive and negative gate conditions, with the verbal WM correlations being of slightly greater magnitude. However, in the initial growth curve models verbal WM was the only statistically significant predictor. Thus, spatial WM also succumbed to multicollinearity and the slightly greater weightings of verbal WM in logic-gate performance. These results suggest that though the factor analysis of WM resulted in two factors, the functional aspect of WM related to learning in the logic-gate training task is unitary and weighted towards the verbal domain.

Training group was the most influential and only variable to consistently correlate with both the true slope and true intercept across all conditions. Training group was also a significant predictor in all of the growth curve models. Consequently, training group was maintained as a predictor in all of the final models, as was verbal WM. PS was represented by verbal speed level in the positive gate models and by spatial speed level in the negative gate models. Finally, spatial WM was dropped as a predictor from all of the models.

The predictor coefficients, R^2 estimates for the true slopes and intercepts, and overall model-fit statistics for the final models are presented in Table 4. The coefficients are in rows 1 through 8, the R^2 estimates in rows 9 and 10, and the model-fit statistics in rows 11 through 15. Two sets of models were evaluated, one that estimated the intercept as initial performance level and a second estimating the intercept as final performance level. Coefficients and R^2 estimates for both sets of intercepts are provided. However, fitting the intercept as endpoint, did not affect either the overall model fit, slope estimates, or the slope/predictor relationships. Accordingly, these estimates are not repeated.

As in the baseline models, the discrepancies among the minimum-fit and Scaled chi-square statistics and the GFI and CFI model-fit indicators show the deviation from multivariate normality. The Scaled chi-square for the negative-gate model was small for its degrees of freedom, $\mathbf{X^2}$ (56) = 69.49, and the Scaled chi-square statistic for the positive-gate model was only slightly large, $\mathbf{X^2}$ (56) = 77.85. The CFI were large for both the positive and negative gate models, CFI = .962 and CFI = .978, respectively, and the SRMR were small, SRMR = .036 and SRMR = .032. Altogether, this shows that the log-

log linear model and the hypothesized relationships of capacity, speed, and training group combined to produce an adequate account of the overall covariance structure.

Table 4 ***Log-Log Linear Logic-Gate Training Models with Predictors***

	Maximum Likelihood Estimates			
Predictor/Parameter	Initial Pos.	Final Pos.	Initial Neg.	Final Neg.
1. VB PSL/INT	.0397	.0815***	--	--
2. VB PSL/SLP	.0201**	--	--	--
3. SP PSL/INT	--	--	.0484*	.0722***
4. SP PSL/SLP	--	--	.0115	
5. VB WM/INT	-.1025***	-.0652**	-.0762***	-.0680**
6. VB WM/SLP	.0179*	--	.0040	--
7. Group/INT	.0861***	.1635***	.0724***	.1360***
8. Group/SLP	.0372***	--	.0306***	--
9. R^2/Int	.233***	.406***	.251***	.394***
10. R^2/SLP	.299***	--	.250***	--
11. (df) χ^2	(56)115.95	--	(56)91.01	--
12. Scaled χ^2	(56)77.85	--	(56)69.49	--
13. GFI	.859	--	.891	--
14. CFI	.962	--	.978	--
15. SRMR	.036	--	.032	--

Note: $^{\dagger}p < .10$; $*p < .05$; $**p < .01$; $***p < .001$; N=119; VB= Verbal; SP = Spatial; PSL = Procedural Speed; WM = Working Memory; INT = Intercept; SLP = Slope.

All of the coefficient estimates were in their predicted directions and all of the coefficients except the VB PS\ INT coefficient in the initial positive-gate model and the SP PS\SLP and VB WM\SLP coefficients in the initial negative gate model were statistically significant (*p*'s < .05). That the PS\INT coefficient in the initial model was not statistically significant and became so in the later model is consistent with the hypothesis of PS having less effect early in skill and increasing over the course of learning. However, that the initial WM\SLP coefficient in the negative gate condition was not statistically significant is at least somewhat enigmatic. One would think that individual differences in WM capacity would be more influential in the more difficult negative-gate condition. One likely explanation is that performance in the negative-gate condition became capacity-limited for some participants, which censored the lower end of the WM distribution and attenuated the effect. The initial negative-gate WM/INT coefficient, which was statistically significant, was also attenuated relative to the positive-gate condition, giving additional support to the hypothesis that the WM effect was attenuated due to capacity limitations.

Interestingly, the negative-gate PS/INT coefficient was enhanced relative to the positive-gate condition, suggesting that procedural speed becomes more influential in performance levels as the effect of WM becomes limited. However, the lack of a significant PS\SLP predictor relationship in the negative gate condition is somewhat puzzling. The simple spatial PS\SLP correlation was statistically significant (see, Table 11). For the predictor relationship in the simultaneous regression not to be significant suggests that somehow the simultaneous regression with verbal WM and training group attenuated the relationship.

Altogether, the predictor effects showed that the participants with greater working memory capacity, who were faster in terms of procedural speed, and who received single-gate training were more efficient in computing digital logic gates, and even more important, they learned at a faster rate. Across the positive- and negative-gate models the combined predictors accounted for approximately 25% of the variance in initial performance levels and rates of change. Over the course of training the predictors came to account for approximately 40% of the variance in performance levels. Thus, performance levels became more predictable with training. Also over training, the proportion of variance accounted for by procedural memory and training method increased, whereas the proportion of variance accounted for by working memory decreased. Altogether, the pattern of relationships among the predictors and the training data are consistent with most theoretical accounts of skill learning and support the hypothesized relationship of repetition priming and procedural memory processes.

Finally, by sequentially entering the predictors into the models ordered by R^2 it was possible to obtain estimates of Cohen's (1988) f^2 effect size (ES) indicator for each predictor. The ES indices for each of the predictor relationships in the logic-gate training models are presented in Table 5. Although the labels are descriptive and approximate, Cohen (1988) characterized f^2 estimates of .02, .15, and .35 as being small, medium, and large relative to the effects commonly observed in psychological research. As predicted, the effect of WM on true intercept decreased from initial performance to final performance. In both the positive and negative gate models the WM ES was effectively halved, decreasing from a medium ES to a small ES. Also as predicted, the effect of PS on true intercept increased from initial to final performance. The PS ES increased seven-fold in the positive gate models and more than doubled in the negative gate models. Though the PS ES estimates continued to be small, the magnitude of the increases was large. Overall, training group had the largest and most cumulative effect on performance levels with participants in the single-gate training group performing both initially and increasingly faster. Training group was also the most influential variable affecting true slopes. It should be noted that on the fourth day of this study, which is not dealt with in this paper, the two training groups performed equivalently in an environment that was identical to the whole-training environment, that is, in terms of mean performance levels. However, individual differences analyses suggested that performance in the single-gate training group was less dependent on working memory capacity.

Table 5 ***Predictor Effect-Size Estimates***

	Positive Gates		Negative Gates	
Parameter	Initial Level	Final Level	Initial Level	Final Level
1. VB PS/INT	.02	.14	--	--
2. VB PS/SLP	.09	--	--	--
3. SP PS/INT	--	--	.05	.12
4. SP PS/SLP	--	--	.04	--
5. VB WM/INT	.17	.08	.14	.09
6. VB WM/SLP	.06	--	.00	--
7. Group/Int	.11	.47	.14	.43
8. Group/SLP	.28	--	.29	--

Note: PS = Procedural Speed; WM = Working Memory; INT = Intercept; SLP = Slope.

Conclusions

Overall, the results of the latent growth curve models supported the hypothesized account of the dynamics of cognitive skill acquisition. As predicted, the models showed initially high working memory loadings followed by decreased working memory and increased procedural memory loadings. Additionally, the models showed that repetition priming is a multi-episodic phenomenon that takes the log-log linear form, and is systematically related to developing skill in a manner consistent with its role as an indicator of procedural memory processing. The models also showed that while the structure obtained in the working memory factor analyses was a domain-specific correlated two-factor solution, the functional aspect of working memory in skill learning is a unitary factor. This is consistent with contemporary accounts of working memory capacity that claim it is a unitary construct with characteristics similar to the psychometric construct of general fluid intelligence (e.g. Engle, Kane, & Tuholski, 1999; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen, & Christal, 1990; Law, Morrin, & Pellegrino, 1995).

It is hoped that the use of these models demonstrates the utility of assessing change, especially in evaluations of learning, and will encourage other researchers to include similar evaluations in their research. Simultaneously assessing the change function and individual differences predictors of change provides a wealth of information not readily available in two-wave data analyses or standard means differences analyses. It allows for an evaluation of the convergence of a complex pattern of results, providing a more thorough test of theoretical predictions. Finally, though this methodology was not conceived of at the time, and in many ways is still in development, I hope this research exemplifies the kind of study envisioned by Cronbach (1957) in his historic and yet to be realized call for the integration of the two disciplines of scientific psychology.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102,* 3-27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117,* 288-318.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Arminger, G., & Schoenberg, R. (1989). Pseudo maximum likelihood estimation and a test for misspecification in mean and covariance structure models. *Psychometrica,54,* 409-425.

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models*. Newbury Park, CA: Sage.

Chou, C., & Bentler P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts Issues, and applications* (pp. 37-55). Thousand Oaks, CA: Sage.

Chou, C., Bentler, P. M., Pentz, M. A. (1998). Comparisons of two statistical approaches to study growth curves: The multilevel model and the latent curve analysis. *Structural Equation Modeling, 5,* 247-266.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (pp. 83-103). New York: Guilford Press.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210*, 207-210.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12,* 671-684.

Cronbach, L. J., & Furby, L. (1970). How we should measure “change” – or should we? *Psychological Bulletin, 74,* 68-80.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.) *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage Publications.

Hu, L. Bentler, P. M., & Kano, Y. (1992). Can statistics in covariance structure analysis be trusted? *Psychological Bulletin, 351-362.*

Engle, R. W., Kane, M. J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). New York: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 3,* 309-331.

Law, D. J. (2000). Working memory and procedural memory: Capacity and speed, separable dimensions in the development of cognitive skill. (Doctoral dissertation, Vanderbilt University, 2000). *Dissertation Abstracts International, 61,* (04), 2234B. (University Microfilms No. AAT 9970058)

Law, D. J., Morrin, K. A., & Pellegrino, J. W. (1995). Training effects and working memory contributions to skill acquisition in a complex coordination task. *Learning and Individual Differences, 7,* 207-234.

Logan, G. D. (1988). Towards an instance theory of automatization. *Psychological Review, 95*, 492-527.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58,* 110-133.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Pellegrino, J. W., & Glaser, R. (1979). Cognitive correlates and components in the analysis of individual differences. *Intelligence, 3,* 187-214.

Poldrack, R. A., Selco S., L. Field, J. E., & Cohen, N. (1999). The relationship between skill learning and repetition priming: experimental and computational analyses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 208-235.

Schacter, D. L., Cooper, L. A., Delaney, S. M., Peterson, M. A., & Tharan, M. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 3-19.

Weinberg, D. R., Diller, L., Gerstman, L., & Schulman, P. (1972). Digit span in right and left hemiplegics. *Journal of Clinical Psychology, 28,* 361.

West, S. G., Finch, J. F., & Curran, P. J. (1995) Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oaks, CA: Sage.

Willett, J . B. (1988). Questions and answers in the measurement of change. *Review of Research in Education, (*Vol. 15, pp. 345-422). Washington, DC: AERA.

Willett, J. B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational and Psychological Measurement, 49,* 587-602.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116,* 363-381.